Opusc. Zool. Budapest, XIV, 1–2, 1977

The Study of Intonation Structure of Bird Vocalizations: an Inadequate Application of Sound Spectrography[1]

By

P. Szőke and M. Filip*

Sound spectrography, as a powerful tool in acoustics in genaral, has also found a wide field of application in ornithoacoustics and given rise to a great body of literature all over the world. In this paper we wish to advance the opinion that the problem of its methodological adequacy in certain applications is ripe enough to be re-examined somewhat more deeply and confronted with another methodological approach as far as the study of *intonation structure* of bird vocalizations is concerned.

I

In linear systems the frequency bandwidth Δf of an oscillatory process is inversely proportional to its duration Δt. This may be expressed in the form of the well known uncertainty principle

$$\Delta f \Delta t \geqq \mu$$

where the dimensionless constant μ depends on how the bandwidth and duration are defined (in nontrivial cases) (Kharkevich, 1962; Goldman, 1948; Küpfmüller, 1949; Winckel, 1960; Pimonov, 1962; Stewart, 1931; Kock, 1935; Gabor, 1946, 1950; Corliss, 1962; Filip, 1970). Applied to acoustics, and verbally interpreted, this means that the two variables, bandwidth and duration, are not mutually independent, that is, tones of limited duration cannot be expected to have an infinitely narrow („line") frequency spectrum: when the duration of a tone decreases its frequency uncertainty increases, and vice versa, if the frequency of a tone is to be determined more definitely the tone should be longer. For a special case of a gated sinusoid the constant μ equals 2.

**Dr. Péter Szőke*, ELTE Állatrendszertani és Ökológiai Tanszék (Zoosystematical and Ecological Institute of the Loránd Eötvös University), Budapest, VIII. Puskin u. 3. – *Dr. Miroslav Filip*, Department of Musicology, Comenius University, Bratislava, Czechoslovakia (†1973).

[1] Fundamental ideas were presented to the 15th International Ornithological Congress, The Hague, August 31 1970 (*Szőke—Tarnóczy—Filip*). Figures of this paper with few exceptions were also projected in the form of slides accompanied by corresponding tape recordings. The projection was repeated and discussed also on the 622nd Session oft he Zoological Section of the Hungarian Biological Society, Budapest, February 5, 1971.

Then, for example, a tone with a 200-msec duration cannot be said to have its „nominal" frequency as a *unique* value since it follows from the uncertainty principle that its frequency-domain representation has a bandwidth $\Delta f = 10$ Hz.

All conventional spectrum analysis methods are linear transforms. This implies that the analysis bandwidth cannot be made arbitrarily narrow in order to attain a high frequency resolution or, in effect, high accuracy if frequency measurement (see above, and: SOANES, 1952; BLACKMAN & TUKEY, 1959; STOPSKIJ, 1962; TOVE & PELC, 1964; KRIKSUNOV, 1965). In other words, the frequency resolution of a filter-type analyzer is again inversely proportional to its time resolution. The more rapid successions of tones are to be analyzed, the wider the analyzing filter bandwidth must be. Consequently, such an analyzer is well suited to measurements where it is the *spectrum density* (or SPL density) that is of primary interest. If, however, the goals of investigation call for measuring the varying *instantaneous frequency* of consecutive tones as a function of time, then conventional spectrography (e. g. the sonogram) is far from being an optimum method.

The most widely known spectrographic method, represented by the "Sonagraph", has been designed specifically for analysis of *formant structure* in human speech.[2] It is quite natural that the instrument has found its many applications in ornithoacoustics, too. For *intonation* studies, however, different methods had been introduced in phonetics and in musicology as early as in 1937 (GRÜTZMACHER & LOTTERMOSER, 1937, 1938, 1940; OBATA & KOBAYASHI, 1937, 1938, 1940).

No doubt, there exist acoustic phenomena and parameters that require the use of a spectrum analyzer, e. g. the Sonagraph, in bioacoustics. On the other hand – and this motivated our paper – bird vocalizations usually have a form of a sequence of more or less distinct tones with more or less definite frequency, including patterns with continuously changing frequency too.[3] If the acoustic structure of bird vocalization is of *this* nature, and if this is what one wishes to study, then a method of *instantaneous frequency* recording is the only adequate approach.[4] The application of the Sonagraph must be viewed just as a routine use of equipment at the scientist's disposal, but is *not* an optimum solution from the methodological point of view.[5]

While spectrography is a frequency-domain method based on the Fourier transform of input time function, the instantaneous frequency (period) measurement is based on the time-domain definition of periodicity (KNESER, 1948; UNGEHEUER, 1963; RIGHINI, 1964; TOVE & PELC, 1964; MCKINNEY, 1965; KORN, 1968; FILIP, 1969, 1970b; LÉON & MARTIN, 1970). The (fundamental) period of a quasi-periodic input signal is then defined by two consecutive posi-

[2] The "sonograph" and "visible speech" techniques began their large bibliography with a paper by R. K. *Potter*, 1945, Science, 102: 463–470, and particularly by a series of papers, 1946, in J. Acoust. Soc. Amer. 18: 1–75, and the book *Potter, Kopp, Green, Visible speech*, New York, Van Nostrand, 1947. An updated bibliography would exceed the scope of this paper.

[3] Sequences of rapid short sounds or crowded micro-patterns often treated in the literature or shown on the sonograms as "impure" or "obscure" or "smeared" sounds, i. e. practically as "noises", may prove to be relatively clear intonations well representable graphically (sometimes also musically) as well as with an instantaneous-frequency recorder, if slowed down sufficiently and professionally.

[4] To the best of our knowledge, the application of instantaneous frequency recording to bird vocalization, has been described in *Fish*, 1953; *Tove—Norman—Isaksson—Czekajeswski*, 1966; *Szőke—Gunn—Filip*, 1969; *Hjorth*, 1970; *Szőke—Tarnóczy—Filip*, 1970.

[5] The frequent statement that the Sonagraph performs the "frequency/time" analysis (*Thorpe*, 1961; *Marler*, 1969; *Thielcke*, 1966; *Hinde*, 1969; *Borror & Halafoff*, 1969 etc.) is to be understood in the sense that *the result of the analysis is plotted in a coordinate system with a frequency axis and a time axis*. The result of the analysis *itself*, i. e. the sonogram, however, is, strictly speaking, a "spectrum-density/time" graphic representation, not a "frequency/time" one.

tive-going (or negative-going) zerocrossings of a signal derived from the input signal in such a way that their steady-state frequencies would be equal. The instantaneous frequency is thus the reciprocal of the instantaneous period and can be displayed by any kind of time-to-voltage converter, in this case perhaps more appropriately termed period-to-voltage converter (Filip, 1970*b*).

In contrast to the Fourier-type analysis, this kind of processing is no longer a linear one and, consequently, is not subjected to the uncertainty principle in the above sense, the accuracy of frequency display being dependent only upon the accuracy of the instrument itself, and upon the signal-to-noise ratio, already for the second period after a transient. Only the "first"-period indication is not valid but this, in fact, is not defined at all as far as real-time processing is taken into account.

The Sonagraph offers the choice of one of the two bandwiths, "wide" and "narrow", usually 300 Hz and 45 Hz. The corresponding time resolution would be inversely proportional, as stated before and, if we take the transient time of a filter to be reciprocal of its bandwidth, we have transient times about 3.3 msec and 22 msec, respectively. It is obvious that the narrow-band analysis with its 22-msec buildup and decay time of filter response would be usable only in the relatively few cases of sufficiently "steady" signals without rapid tone successions,[6] and even the 45-Hz band would represent, for example, about a third-ovtave "uncertainty" with respect to a centre frequency of 200 Hz. Moreover, the narrow-band analysis with its considerably decreased time resolution could hardly be used in analysis of frequency modulated tones with high rates of frequency modulation (Marler, 1969) which is found often in bird vocalizations.

Thus, the question is not which of the two bandwidths is more appropriate in general, but *whether sound spectrography is an appropriate method at all.* It was our aim to show that it is not, if we wish to study the *intonation structure* of bird vocalizations. Then the instantaneous frequency recording is the only methodologically justified approach to the problem.[7]

In instantaneous-frequency recording it is conceptually useful to discern two stages of signal processing, namely, the *extraction* stage (extraction of periodicity information from the complex input signal) and *conversion* stage (period-to-voltage conversion and recording). The instantaneous frequency graphs presented in this paper have been obtained with an instrument developed by the second author for the Institute of Musicology, Slovak Academy of Sciences, Bratislava (1961 – 64). The extraction method which may be termed envelope periodicity detection has been described elsewhere (Filip, 1969), so we will not give the details. The conversion method has been modified from that of Grützmacher and Lottermoser, well known in the literature (Grützmacher & Lottermoser 1937, 1938, 1940; Filip, 1970, 1970*b*).

[6] The well known phenomenon of "ringing" is one of the manifestations of limited time resolution. To achieve sufficient frequency resolution the selectivity of the filter must be sufficiently high. Then its impulse response (inverse Fourier transform of its transmission function) is not aperiodic and the duration of "ringing" is, roughly speaking, proportional to the selectivity. When a very short tone burts is being applied, the filter responds by several periods of its resonance frequency thus effectively prolonging the apparent duration of the measured tone. (See also *Davis*, 1964, p. 127.) The delayed ringing of the narrow-band filter, being even longer than the sound itself, often makes the thin vertical transient lines thick (as for example in Figs 9*N*, 11*N*).

[7] It has been used by the authors of this paper since 1964 and reported on various oceassions including the 14th Int. Ornithol. Cong., Oxford 1966, as well as at the same time in an informal meeting and unpublished communication.

As concerns the relation between objective and subjective methods of representation (on the basis of sound microscopy, see Section II), it is hoped that it is worth mentioning that the periodicity pitch perception in quasiharmonic signals such as musical tones and vowels has a much closer analogy in the time-domain definition of frequency as implemented in instruments for recording the instantaneous-frequency graphs (NORDMARK, 1968; FILIP, 1970*a*) than in the spectrum-density definition implemented by pure FOURIER-type analyzers. Thus the high correlation between objective and subjective graphs has indeed more than just a practical value.

It should also be pointed out that for the (human) ear, as a nonlinear system, the uncertainty principle no longer holds in its original sense, valid for linear systems (LIANG & CHISTOVICH, 1960; CARDOZO, 1962; SEKEY, 1962, 1963; SCHIEF, 1963; MAJERNIK, 1964, 1967; KURZE, 1965; GAMBARDELLA & TRAUTTEUR, 1966; CORLISS, 1967; RONKEN, 1971) and thus the time resolution ability is much greater than it would be if it were determined by the uncertainty principle for linear systems and by the admirably high frequency resolution of human auditory system. It is assumed that birds' frequency resolution is comparable to that of man (KNECHT, 1940; SCHWARTZKOPFF, 1949, 1952, 1955; GALAMBOS, 1954; THORPE, 1961[8]). There is some evidence, however, that their *time resolution* capabilities are higher than in man (MARLER, 1969) and important experiments have been described which indicate that it is *considerably higher* (KONISHI, 1969[9]). The multiple slowing down of bird vocalization tape recordings when studied and notated by ear may thus be seen as a compensation for the difference in the time resolution properties of human and avian auditory system.

If we again take the reciprocal of frequency bandwidth as the transient time of a filter, and consider this time to be representative for the time resolution of the analysis, then the time resolution of the instantaneous-frequency measurement (i. e., one cycle) may be shown to be *Q times better* than that of the spectrograph, where Q is the quality factor of the (idealized LC) filter, $Q = f_0/\Delta f$, with f_0 being the centre frequence of the filter and Δf its half-power-point bandwidth.

Moreover, the frequency records represent the instantaneous frequency in the form of a line (or, equivalently, in the form of a boundary between black and white areas as in this paper) with practically *negligible* width, in accordance with the time-domain definition of instantaneous frequency as opposed to the spectrum-density definition implemented by a spectrum analyzer. Thus the *frequency resolution* of instantaneous-frequency measurement is practically equal to the measurement accuracy which is, as stated above, limited only by the properties of the equipment and by the signal-to-noise ratio. This is indeed negligible if compared to the width of sonogram traces.

[8] Some direct quotations from *Thorpe*, 1961: "It is everywhere agreed that frequency-analysis or harmonict analysis is the essential basis of 'hearing' in at least higher vertebrates — that is to say, the fish, birds and mammals — as against hearing by the analysis of amplitude-modulation which predominates in the insects" (120). . "The presenl overall picture of the hearing abilities of birds which thus emerges suggests that it is similar to our own in genera-range and ability to discriminate pitch. Song birds and parrots certainly approach human abilities. . ." (127). . . "In conclusion we can say, with *Galambos*, that the capacity for dealing with tones, as measured by psychological testing, is not remarkably dissimilar for fish, birds and men" (128).

[9] "Most units (avian auditory neurones) exhibit near 100 per cent time-locking to a train of clicks when the inter-click interval exceeds 1.3 – 2.0 ms". . . Units can follow click repetition rates lower or higher than their best frequencies (to which the units are most sensitive), although few units can follow on a one-to-one basis repetition rate higer than about 1000 clicks per s. . . "In comparison with songbirds, a species like the pigeon does not seem to have any rapid sequence of sounds in its vocalizations, yet its auditory neurones can resolve such sounds" (pp. 566 – 567).

Besides, our instantaneous-frequency graphs (see Figs 3*F*, 4*F*, 5*F*, 6*F*, 15*F*) are also intended to confirm the validity, in its specific sense, of both graphic or semigraphic (see e. g. Figs 10*G*, 7*S*) and adapted musical notation in portraying musically structured bird vocalizations, the graphic and semigraphic representation being suitable also for those structured non-musically.

II

As a result of his developing of *sound microscopy* (the prerequisite of maximum possible adequacy of any *subjective* representational mode) into a *fundamental and consistently applied research method*, the first author has realized and comprehensively examined the apparent consequences of the inadequacy of sound spectrogrpahy in its present-day ornithoacoustical applications, and due to the facts revealed, a (1) *graphic* or (2) *semigraphic* (five-line staff) representational mode and an (3) adapted *musical notation* have been developed by him for the purposes of the sufficiently reliable aural (subjective) transcription of the intonation (pitch and time) structure of bird vocalizations (similarly, in some sense, to that known in ethnomusicology, for example).[10] This new method of musical representation based on sound microscopy and applicable only for musically structured vocalizations is, of course, basically different from the earlier dilettante and naive attempts of "musical transcription" of natural bird sounds, applied even to those structured non-musically.

As the *conventional* musical notation is a compund graphic and *symbolic* representation, it is fully satisfactory only in its original application, i. e., roughly speaking, to professional (composed) music except "New Music". In ethnomusicology, as well as in the study of "musically" structured avian vocalizations, the transcription calls for certain refinements and additional signs to complement the traditional ones, and at the same time for the reexamination and clarification of some traditional views and concepts of theoretical and practical importance concerning music in genaral and avian musicality in particular[11] (Hartshorne, 1958; Thorpe, 1961; Thorpe & Lade 1961*a*; Davis & irby, 1964; Halafoff, 1968; Hinde, 1969; Hold, 1970).

The purely graphic representation, as already mentioned, may also be used to represent acoustical phenomena not expressible by conventional, or even adapted musical notation. In fact, the graphic (and semigraphic) representation is a subjective analogy to the objective instantaneous-frequency records and can be regarded, in a sense, as if the physical record were subjected to a kind of data reduction process carried out by the pitch and time perception mechanisms. Thus, from the physical signal represented by the frequency graph, and sub-

[10] "Subjective" is not to be confused with arbitiary" or "biased". All three subjective (graphic, semigraphic and musical) modes of representation applied in this study are justified exclusively on the basis of high (in Passeriformes regularly 16—64-fold, rarely 128-fold) stretch of time (scientifically demanding slowing down the speed of vocalization).

[11] In order to avoid misunderstandings, we have to explain, though with some simplification here, at least two basic terms in this paper. As *musical* are treated bird sound phenomena (intonation structures) based on a *tonal system* (scale) and consisting mostly of tones with "musical pitch invervals" known to us from human music (first of all from folk music) and, on the physical level, analogue to the harmonic (or possibly quasiharmonic) frequency intervals (relations of overtones in so far they are discriminable and learnable by the avian (and human) hearing. Acoustic phenomena without "musical" intervals in this sense are concerned as *non-musical*. In the last analysis, "musicality" in birds as well as in man (and even in the pure physical inorganic sphere) is, in the light of the facts, a question of specific pitch (frequency) structure, not of function or meaning. Consequently, the "musical" character of bird vocalizations is also independent of their time (rhythmic or non-rhythmic) structure although the process of rhythmlzation (repetition) had an important share in the evolution of different forms of "avian music" (and, equally, of the developed non-musical forms of bird vocallzation) (*Szőke*, 1974).

jected to the sound microscopy technique, a perceptual pattern results which is written down by an experienced scholar in the form of a graphic representation. In order to make an immediate auditory (pitch) imagination possible, the semigraphic notation combines the features of pure graphic representation with the advantages of musical five-line staff (in the case of both non-musical and musical structure of vocalization).[12]

Professionally made subjective transcriptions based on the necessary sound microscopy can be regarded as close and practically sufficient approximations to the psychoacoustical (perceptual) form of bird vocalizations in their pitch and time aspects. The accuracy of the subjective modes of representation (1, 2, 3) presented here can be still more refined to a reasonable measure although even at their present state they give us much more and more precise information about the pitch and time structure of bird vocalization than the objective sonography inadequately applied (see, for example, Figs 7, 8, 9, 11 with the corresponding text and SZŐKE – GUNN – FILIP, 1969).

In the following, various examples of bird vocalizations are represented by different methods in order to verify *in practice* the theoretical statements and ideas expounded so far. On some of the sonograms the traces of higher harmonics were eliminated (covered) so that they do not interfere in confrontation with other representations revealing only the fundamental frequency.

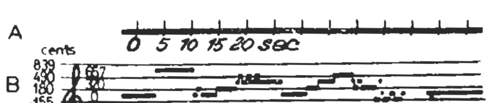
III

On the following pages and in the Figures some abbreviations will be used, namely: *W* for "wide band", *N* for "narrow band" sonogram, *F* for "fundamental (instantaneous) frequency graph", *G* for "graphic", *S* for "semigraphic" and *M* for "musical" representation.[13]

Fig. 1. The *W* filter smears the frequency "lines" as though they were being painted with a wide brush (DAVIS 1964), so that it is not possible to read out the instantaneous (pitch-) frequency. This is also symbolized in the musical five-line staff with note heads unreadably large.

Fig. 2. Musical semitone scale descending from C_4 to C_3 (with rounded off frequency values in *G*) played by the first author on a wind-instrument. The

[12] Of similar "semigraphic" character are the experimental folk music notations made by means of a computer in the Royal Institute of Technology, Stockholm. An example reproduced from *Sundberg & Tjerlund*, 1971:



[13] *Explanations of signs used in the Figures: Time data* on the left (e. g., 1.4 sec, 0.7 s etc.) refer to the natural, i. e. not slowed down duration of the vocalization illustrated. *Numerals with arrows:* 2 ↑, 3 ↑, 4 ↑ above the clef mean that the natural pitch is 2, 3 or 4 octaves higher, respectively, than notated, while 1 ↓, 2 ↓, 3 ↓ below the clef mean that the pitch of the slowed down reproduction is 1, 2 or 3 octaves lower, respectively, than notated. *Prolongation or shortening of notes:* ⌒ means a slight extension, while ◡ a slight shortening of value of the marked note. *Numbers in squares:* 16, 32, 64 mean that the natural duration is stretched (i. e. the speed slowed down) 16, 32 or 64 times, respectively. *Metronome marking,* for example = 60, indicates the approximative tempo of the slowed down vocalization. i. e., in this example 60 quarter notes per minute. (Subjective illustrations in some Figures give data in "centiseconds". To obtain standardized indication in milliseconds please multiply the given numbers by 10.)

jected to the sound microscopy technique, a perceptual pattern results which is written down by an experienced scholar in the form of a graphic representation. In order to make an immediate auditory (pitch) imagination possible, the semigraphic notation combines the features of pure graphic representation with the advantages of musical five-line staff (in the case of both non-musical and musical structure of vocalization).[12]

Professionally made subjective transcriptions based on the necessary sound microscopy can be regarded as close and practically sufficient approximations to the psychoacoustical (perceptual) form of bird vocalizations in their pitch and time aspects. The accuracy of the subjective modes of representation (1, 2, 3) presented here can be still more refined to a reasonable measure although even at their present state they give us much more and more precise information about the pitch and time structure of bird vocalization than the objective sonography inadequately applied (see, for example, Figs 7, 8, 9, 11 with the corresponding text and Szőke – Gunn – Filip, 1969).

In the following, various examples of bird vocalizations are represented by different methods in order to verify *in practice* the theoretical statements and ideas expounded so far. On some of the sonograms the traces of higher harmonics were eliminated (covered) so that they do not interfere in confrontation with other representations revealing only the fundamental frequency.

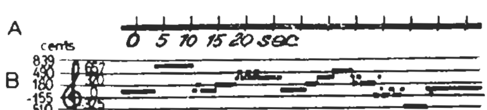
III

On the following pages and in the Figures some abbreviations will be used, namely: *W* for "wide band", *N* for "narrow band" sonogram, *F* for "fundamental (instantaneous) frequency graph", *G* for "graphic", *S* for "semigraphic" and *M* for "musical" representation.[13]

Fig. 1. The *W* filter smears the frequency "lines" as though they were being painted with a wide brush (Davis 1964), so that it is not possible to read out the instantaneous (pitch-) frequency. This is also symbolized in the musical five-line staff with note heads unreadably large.

Fig. 2. Musical semitone scale descending from C_4 to C_3 (with rounded off frequency values in *G*) played by the first author on a wind-instrument. The

[12] Of similar "semigraphic" character are the experimental folk music notations made by means of a computer in the Royal Institute of Technology, Stockholm. An example reproduced from *Sundberg & Tjerlund*, 1971:



[13] *Explanations of signs used in the Figures: Time data* on the left (e. g., 1.4 sec, 0.7 s etc.) refer to the natural, i. e. not slowed down duration of the vocalization illustrated. *Numerals with arrows:* 2 ↑, 3 ↑, 4 ↑ above the clef mean that the natural pitch is 2, 3 or 4 octaves higher, respectively, than notated, while 1 ↓, 2 ↓, 3 ↓ below the clef mean that the pitch of the slowed down reproduction is 1, 2 or 3 octaves lower, respectively, than notated. *Prolongation or shortening of notes:* ∩ means a slight extension, while ∪ a slight shortening of value of the marked note. *Numbers in squares:* $\boxed{16,}$ $\boxed{32,}$ $\boxed{64}$ mean that the natural duration is stretched (i. e. the speed slowed down) 16, 32 or 64 times, respectively. *Metronome marking*, for example = 60, indicates the approximative tempo of the slowed down vocalization. i. e., in this example 60 quarter notes per minute. (Subjective illustrations in some Figures give data in "centiseconds". To obtain standardized indication in milliseconds please multiply the given numbers by 10.)

overlapping of *consecutive* tones is clearly seen in N but not, with one exception, in W. The overlapping is caused by *ringing* (i. e., decay transient of the N filter). In N the only tone F_3 in the middle of the scale, intentionally played shorter than the others, seems almost to touch the subsequent one: here the ringing effect covers almost the whole rest between the two tones. The real endings of the tones are marked by pale vertical transients (see Footnote 6). This clear non-avian example makes easier to understand the artifacts that may be caused by ringing under more complex circumstances and when combined with other effects in birds.

Fig. 3. Repeated musical "horn motifs" as a portion of a song of the Great Tit *(Parus major)*. F with its linear semitone (logarithmic frequency) calibration refers clearly to the musical pitch structure (intonation contour) perceived when slowed down 32 times and illustrated in M. Whereas W masks the frequency (pitch-) structure.

Fig. 4 W, F, M. Grey Warbler *(Gerygone igata)* song (recorded by K. & J. Bigwood, New Zealand) of surprisingly folksong-like three-section form (see in F, M and W) with short "introductory" part (a) and a recitative "rhythm" (b). In contrast to the frequency-smearing effect of W, graph F (without the "introductory" part) displays clearly the fundamental frequency as a function of time. Note the convincing parallelism (analogy) between the objective F and the subjective M. (In W the two initial tones of M are not recorded.)

Fig. 5. Yellow breasted Tit *(Petroica macrocephala)* song of folksong-like one-section form (recorded by K. & J. Bigwood, New Zealand). F (here with logarithmic semitone calibration) refers again to the musical perception of the song structure shown in M on the basis of a 16 and 32 times slowed down playback. In F the pitch level of the song is not quite fixed, and in M is slightly higher than in fact.[14] The time structure can be portrayed still more precisely if represented graphically (or semigraphically).

Fig. 6. Hermit Thrush *(Hylocichla guttata)* song (recorded by D. J. Borror, Ohio). F and M show the musical micro-structure of the second part of the song, set in frame on W.

Fig. 7. The initial part of a song of the Wren *Troglodytes troglodytes*. Erroneously the bird is regerded as one of the most famous avian "musicians" of Europe. Its song, however, is non-musical, for it consists only of slurring (glissando) sounds the continuously changing frequency (intonation contour) of which is also obscured by the large bands in W. In contrast, S (based on 32- and 64-fold time stretch) displays the non-musical perceptual pattern (i. e. the intonation contour involving the time structure) distinctly. The semigraphic mode of representation is utilisable by scientists both unacquainted and familiar with the music-reading. The trained note-readers, however, are able to decode the semigraphic (five-line staff) representation much better (even in the case of non-musical structure of bird vocalization), especially concerning the *sounding* forms of bird vocalization (slowed down).

[14] The general pitch level (tessitura) of singing of birds often undergoes some slight and insignificant continuous changes (temporary small-scale decreasing or incresing of frequency) without to be perceivable on the auditory level of birds (and of man as well). As in general the slight frequency (pitch) level instability of this kind (as well as the simultaneously arising subliminal tenporal inequality of tones of "equal duration", too) escape notice even at a speed slowed down greatly, so, in contrast to the objective frequency graphs, these latent and contingent changes as inessential physical phenomena cannot and must not be visualized in the subjective graphic and musical illustrations. This, however, does not mean that the structural analogy between the physical and the psycho-physiological processes of vocalization and hearing in birds is questionable.

Fig. 8. Redstart *(Phoenicurus phoenicurus)* song of non-musical structure. (In *W* the initial part of the long introductory tone has been accidentally cut off.) Here too the intonation structure remains unrecognizable in contrast to *S* offering a full and reliable picture of the pitch (frequency) and time pattern of the song.

Fig. 9. One of the highly developed micromelodies of the Hermit Thrush *(Hylocichla guttata)* (recorded by W. W. H. GUNN, Canada). When slowed down greatly the detailed pitch and time structure can easily be revealed even by an unexperienced listener as a surprisingly "human-like" song form (*M*). The slow avian melody in *M* was resung by the first author and speeded up again 32 times to be spectrographed in order to compare the new spectrogram W_1 with that of the original avian song *W*. It is no surprise that in W_1 the "human-like" musical character of the tune disappeared completely and to our ears (and eyes) the song structure became unrecognizable again. The long vertical lines in W_1 represent buildup and decay transients. *N* is a narrow band variant of W_1. It also distinguishes itself by long vertical transient lines. However, as a consequence of ringing, in *N* these vertical transient effects grew misleadingly thick (extended rightwards on the sonogram). (The reverbation-like pale multiplication of markings on some narrow band sonograms, as in Figs 9*N*, 18*N*, is an eliminable sort of distortion caused by the Sonagraph.) Further on, the slow melody *M* was played by the first author on a wind-instrument in order to obtain a maximally clear intonation contour, then this was speeded up again to about the natural pitch level and duration of the original bird song, and spectrographed as well (N_1). Comparison of *W* of the natural avian song with N_1 of this man-performed instrumental avian tune shows how in the latter the ringing smears even the separate, not directly neighbouring tones of the same frequency into seemingly continuous long horizontal traces.

Fig. 10. Great Tit (*Parus major*) call structured musically. In *W* the long vertical lines caused by the buildup (onset) and decay transient responses of the Sonagraph accur at sufficient distances from each other not to cause too much disturbance (though here the width of bands also makes the fundamental-frequency structure rather unrecognizable). *G* and *M* show the musical pattern clearly.

Even with wide band spectrograms insurmountable difficulties may arise from short sounds with rapid attacks resulting in (long verical) wide-band transient lines in the record. The more with *narrow band* spectrograms rapid transients and frequency modulated tones (e. g. vibratos) would present time-resolution problems that could hardly be overcome, as already mentioned in Section I.

Fig. 11. Grasshopper Warbler (*Locustella naevia*). Song portion consisting (differently from that in Fig. 10) in a *rapid* succession of *short* discrete tones (about 1500 micro-motifs per min) with a dense series of extremely wide transient responses by the Sonagram (the long vertical lines in *W*) which make the actual vocal pattern, shown clearly in *S* and *M*, totally unrecognizable. In *N* the vertical transient lines are considerably thickened and smeared together by the ringing effect of the narrow-band filter. Thus *N* becomes still more crowded than *W*. (See Footnote 6.)

Fig. 12. Six different fractions *ABCDEF* of a continuous song of the Sky Lark *(Alauda arvensis)*. The comparison of mutually corresponding rhythmic patterns *A* through *F* (in *W* and *M*) shows an unacceptable smearing effect of the bandwidth and the dense and long transient lines produced by the Sonagraph. *W* masks, *M*, however, discloses the rhythmic patterns of the song fractions.

Fig. 13. Due to the dense groups of vertical transient lines and the large bandwidth in *W* it is impossible to recognize the characteristic primitive musical rhythmic pattern of the song of the Grasshopper Sparrow (*Ammodramus savannorum)* (recorded by W. W. H. Gunn, Canada), though in this case *W* was made exceptionally at a speed slowed down 2 times. When stretched 64 times (with the necessary careful technique), the rhythmic pattern is audible distinctly in all details represented in *M*. Note the almost perfect regular alternation of four-beat and five-beat micro-measures. (*W* starts at "s." in *M*.)

Fig. 14. River Warbler *(Locustella fluviatilis)*. Song portion repeating rapidly (about 600 times per min) a longer and more complex musical micro-motif (*S*) of about 7 – 8 msec duration with extremely crowded transient phenomena and wide frequency bands in *W*. The musical structure *M* of the song (with some short non-musical slurrings), although distinctly audible in all details at a tape speed reduced 32 or 64 times, is completely concealed in *W*. This is an expressive example of the different reliability of the representational methods shown in *W*, *S* and *M*.

Fig. 15. Ortolan Bunting *(Emberiza hortulana)* song. To the "trills" and the rapid final warbling (vibrato *V* in *W*) of the song the instantaneous-frequency recorder responds (in *F*) with clear indication of instantaneous pitch-frequency level (though here not sufficiently expressive due to the excessive typographic reduction in size of the original graph), while *W* does it with too wide bands caused by close succession of vertical transient lines smeared together. Compare the four different representations V, V_1, V_2, V_3, of the warbling final tone. In *F* (if V_2 stretched still more than represented here) the rate of warbling, about 350 cycles of fluctuation per second, can easily be calculated due to the adjustability of the speed of film. Of course, independently of this rapidity the warbling can also be made audible in all details when slowed down 64 times.

Special difficulties arise when frequency modulated tones (e. g. vibratos), so common in bird vocalizations, are to be spectrographed with narrow band filter. If the modulation frequency becomes greater than the filter width of the Sonagraph (wide band: usually 300 Hz, narrow band: 45 Hz), then the sonogram no longer shows the vibrato, i. e. the periodic frequency fluctuation of the signal. Instead the signal, i. e. the warbling tone, is split up into several horizontal *side bands* running one above the other. The resolving of the frequency modulated final tone (V, V_2, V_3) into side bands is displayed, for example, in V_1 (Fig. 15).

Fig. 16. Corn Bunting *(Emberiza calandra)* song. Behind the conspicuous vertical transient lines and wide frequency bands in *W* a very complex, predominantly musical, structure is hidden (displayed semigraphically in *S*). This is an expressive example of how the wide band sonogram masks visually the "pitch" and the time structure of complex bird songs. Compare, for instance, the five initial "flag-like" bands in *W* with the corresponding five initial micro-motifs (with corresponding numbering) of musical structure, composed

of several discrete short tones as shown in *S*, displaying distinctly their whole fine structure smeared in *W*. Existence of such well-formed and richly patterned bird songs, designating both the species and the individual, cannot be merely the acoustic result of latent (innate) myogenic or neurogenic processes, but, in the last analysis, it can, in some sense, only result from the social life of the species (of course, on physical and psycho-physiological basis). This means that the pattern (intonation contour) of such complex songs must be audible to the birds in every structural detail.

Fig. 17. The well known call of the Greenfinch *(Carduelis chloris)*. *W* gives a sound picture blurred by the extreme density of long transient lines, presenting striking evidence of heavy visual masking of frequency and time structure of bird calls consisting of rapid successions of discrete short tones (here about 140 per second) with more or less definite pitch, as shown in *G* at a speed slowed down 64 times, representing only a short portion of the call. With its series of discrete apparent transient lines another wide band sonogram W_1 made at a two-fold stretch of time, reveals the periodically interrupted (tremolo, buzz-like) time structure of the call, hidding at the same time the frequency (pitch) structure by the wide (long) transient lines. By counting the apparent transient lines (in fact smearing together both the buildup and the decay effects) it may be found that the mean rate of sound bursts in the tremolo is about 160 per second, decreasing towards the final part of the call to about 145 bursts per sec. This agrees sufficiently with the result 140/s obtained through aural counting of the sound bursts at a tape speed slowed down 64 times (*G*). In *N* the series of vertical transient lines characteristic of such wide band sonograms becomes transformed into an extremely wide band of fluctuating and interlocked long *horizontal* frequency lines lying closely on one another. Here both the time resolution and the frequency resolution are insufficient.

Fig. 18. Short (0.5 s) alarm call of the Great Tit *(Parus major)*, musically structured and containing three warbling tones, i. e. vibratos *(a, b, c* or *1, 2, 3)*. Graphic representation *G* of a short portion of the warbling tone *a (= 1)* shows a warbling of 225 vibrato cycles per second which are audible distinctly if slowed down appropriately. *N* is distorted by side bands of the warbling tones *1, 2, 3* as well as by other smearing effects (e. g. ringing).

Here we have come to the end of the demonstration of our practical examples showing that the actual intonation patterns of bird vocalizations must be analyzed in an adequate way. Without a thorough and demanding knowledge of bird vocalizations, without the knowledge of their pitch (frequency) and time structure, it is difficult, if not impossible, to study certain essential aspects of avian life sufficiently comprehensively and reliably. Moreover, the significance of ornithological acoustics (implying ornithomusicology) for musicology in general, for the disclosure of the pre-human and, in general, presumably biological fundamentals of human music aesthetics, and for some other natural and social sciences, also calls for more adequate methods. These methods imply, of course, the necessity for a deeper study of birds' auditory mechanisms, first of all their time discrimination.

The main aim of this paper is not to demonstrate our experimental new representational methods. However, if confronted with the sonograms conventionally applied, these new methods seem to be suitable making the inadequacy of the sonograms in the applications mentioned evident. Further, we believe that,

in general, these methods (both objective and subjective) represent a promising direction in which the study of intonation structure of bird vocalizations (and in most cases of those of fishes, amphibia, reptilia and mammalia too) may be succesfully developed.

Acknowledgements

The authors are greatly indebted to Professor LAJOS TAMÁS, Member of the Hungarian Academy of Sciences and Director of the Hungarian Institute of Linguistics, for the Sona-Graph (6061 A, Kay Co.) which he kindly put at the authors' disposal at the Phonetic Laboratory of his Institute, as well as to TAMÁS TARNÓCZY, Professor of physical acoustics at the Eötvös Loránd University of Sciences, Budapest, and Director of the Acoustical Research Group of the Hungarian Academy of Sciences, for his criticism of some problems discussed in the paper and for the facilitation of preparing the sound spectrograms presented.

Summary

The spectrographic method producing the well known sonograms has found considerably wide use in bio-acoustics. In certain , particularly in some ornitho-acoustical applications it must be borne in mind, however, that it represents the *spectrum density* as a function of time and frequency. The frequent statement that the Sonagraph performs the "frequency/time" analysis is to be unders- in the sense that the result of the analysis is plotted in a coordinate system with a frequency axis and a time axis. The result of the analysis itself, i. e. the sonogram, is, stricly speaking, a " spectrum-density/time" graphic representation, not a "frequency/time" one. If we study, however that parameter of acoustical stimuli which is perceived as pitch (varying or constant, changing stepwise or continuously) or, in short, if it is the *intonation structure* of avian vocalization that is to be investigated, then the only adequate method is the recording of *instantaneous frequency* defined in the time domain. This has long been recognized in experimental phonetics and in musical acoustics but, with only a few exceptions, not in bio-acoustics. The paper aims at demonstrating that for the study of intonation (pitch and time structure) the spectrographic method is far from optimum because it is a linear method, subjected to the uncertainty principle. Particular difficulties are encountered with rapid successions of short sounds and in frequency modulated or periodically interrupted tones. On the other hand, the time resolution of the instantaneous-frequency measurement is one cycle, and its frequency resolution depends only on the accuracy of the measuring equipment and on the signal-to-noise ratio. The instantaneous-frequency graphs also confirm the validity and scientific value of the graphic, semigraphic and adapted musical representation based on sound microscopy, shown in the paper with corresponding sonograms.

ZUSAMMENFASSUNG

Die Anwendung der mittels der linearen Methode der Klangspektrographie produzierten Sonogramme ist in der Bioakustik allgemein verbreitet. Dabei wird aber die Tatsache außer acht gelassen, daß die Sonogramme nicht die Frequenz der Töne, sondern deren *Spektrumbreite* (Spektrumdichte, spectrum density) in der Funktion der Zeit darstellen. So repräsentiert das Sonogramm nicht den von den Vögeln (und auch vom Menschen) als Tonhöhe wahrgenommenen Parameter

der akustischen Stimuli, sondern einen anderen physikalischen Parameter, der aber im Tonhöhenunterscheidungsvermögen der Vögel keine Rolle spielt. Die bioakustische Literatur behauptet, daß das Sonograph eine „Frequenz/Zeit"-Analyse produziert. In der Tat jedoch analysiert das Sonograph die „Spektrumdichte/Zeit"-Struktur, und das Sonogramm – das sich zwar im „Frequenz/Zeit"-Koordinatensystem ausprägt – veranschaulicht eigentlich die „Spektrumdichte/Zeit"-Struktur, nicht aber die „Frequenz/Zeit"-Struktur der Vogelstimme. Das bedeutet praktisch, daß das Sonogramm die „Intonationsgestalt", d. h. die wirkliche „Tonhöhe/Zeit"-Struktur der Vogelstimmen nicht abbildet, sondern *verdeckt*. In unserem Aufsatze wird bewiesen, daß sich zur Darstellung der Intonationsgestalt die spektrographische Methode im Prinzip nicht eignet. Die Aufgabe daher ist nicht, diese Methode zu vervollständigen, sondern sie mit einer anderen – nicht-linearen, adäquaten und exakten – physikalischen Methode zu ersetzen, bei welcher das Prinzip der Unsicherheitsrelation (uncertainty principle) nicht zur Geltung kommt. Das Unsicherheitsprinzip verursacht besonders bei der Analyse von rapiden Tonsukzessionen und Frequenzmodulationen unüberbrückbare Schwierigkeiten.

Die einzige, prinzipiell adäquate Methode der Untersuchung der „Frequenz/Zeit"-Struktur, d. h. der Intonationsgestalt der Vögelstimmen ist die Darstellung der *momentanen Frequenz*. Diese nicht-lineare objektive Methode bekräftigt auch die psychoakustische Adäquatheit der auf Grund der starken *Tonverlangsamung* sachgemäß hergestellten subjektiven (korrelativen) graphischen und biomusikalischen Abbildungen der Vogelstimmen.

REFERENCES

1. Blackmann, R. B. & Tukey, J. W. (1959): *The measurement of power spectra from the point of view of communication engineering.* – New York, Dover Publ.

2. Borror, D. J. & Halafoff, K. C. (1969): *Notes on song structure in the Townsend's Solitaire.* – Wilson Bulletin, 81: 163–168.

3. Cardozo, B. L. (1962): *Frequency discrimination of the human ear.* – 4th Int. Cong. Acoust., Copenhagen, Paper H16.

4. Corliss, E. L. R. (1962): *Indication limit.* – 4th Int. Cong. Acoust., Copenhagen, Paper N22.

5. Corliss, E. L. R. (1967): *Mechanistic aspects of hearing.* – J. Acoust. Soc. Amer., 41: 1500–1516.

6. Davis, L. Irby (1964): *Biological acoustics and the use of the sound spectrograph.* Southwestern Naturalist, 9 (3): 118–145.

7. Filip, M. (1969): *Envelope periodicity detection.* – J. Acoust. Soc. Amer., 45: 719–732.

8. Filip, M. (1970): *Frekvencné merania a tónová sústava. (Frequency measurement and tone system.* In Slovakian with German summary.) – In: Nové cesty hudby (2): 50–85, Prague.

9. Filip, M. (1970*a*): *Auditory and electronic envelope periodicity detection.* – J. Acoust. Soc. Amer., 47: 654–658.

10. Filip, M. (1970*b*): *A survey of pitch-frequency recording methods.* – Symposium on intonation (Int. Soc. Phonet. Sci.), Prague, in print.

11. Fish, W. R. (1953): *Method for the objective study of bird song and its application to the analysis of Bewick Wren songs.* – The Condor, 55: 250–257.

12. Gabor, D. (1946): *Theory of communication* – Journ. IEE 93, Pt. III: 429–457.

13. Gabor, D. (1950): *Communication theory and physics.* – Philos. Magaz., 41, Ser. 7: 1161–1187.

14. Galambos, R. (1954): *The neural mechanisms of audition.* – Physiol Rev., 34: 495–528 (23, 37).

15. Gambardella, G. & Trautteur, G. (1966): *Time-frequency analysis in the hearing process.* J. Acoust. Soc. Amer., 40: 1187–1189.

16. Goldman, S. (1948): *Frequency analysis, modulation and noise.* New York, McGraw-Hill.

17. Grützmacher, *M.* & Lottermoser, *W.* (1937): *Über ein Verfahren zur trägheitsfreien Aufzeichnung von Melodiekurven.* – Akust. Z., 2: 242–248.

18. Grützmacher, M. & Lottermoser, W. (1938): *Die Verwendung des Tonhöhenschreibers bei mathematischen, phonetischen und musikalischen Aufgaben.* – Akust. Z., 3: 183–196.

19. Grützmacher, M. & Lottermoser, W. (1940): *Die Aufzeichnung kleiner Tonhöhenschwankungen.* – Akust. Z., 5: 1–6.

20. Halafoff, K. C. (1968): *A survey of birds' music.* – Emu, 68, Pt. 1, June: 21–40.

21. Hansen, P. (1974): *The sonagraphic interpretation of glissando figures in bird vocaltzations.* – Biophon, II, 3: 3–6. (International Bioacustics Council, Aarhus).

22. Hartshorne, C. (1958): *The relation of bird song to music.* – The Ibis, 100: 421–445.

23. Hinde, R. A. (1969): *The physical analysis of avian vocalizations. Introduction.* – In: Bird Vocalizations. Cambridge University Press: 1–3.

24. Hjorth, I. (1970): *A comment on graphic displays of bird sound and analyses with a new device, the melograph Mona.* – J. Theoret. Biol., 26: 1–10.

25. Hold, T. (1970): *The notation of bird song: a review and recommendation.* – The Ibis, 112: 111–172.

26. Khárkevich, A. A. (1962): *Spektry i analiz.* – 4th ed. Moscow, Fizmatgiz (1st ed., 1957). English translation: Kharkevich A. A. 1960. Spectra and Analysis. New York, Consultants Bureau.

27. Knecht, S. (1940): *Über den Gehörsinn und die Musikalität der Vögel.* – Zeitschr. f. vergl. Physiol., 27: 171–232 (126).

28. Kneser, H. (1948): *Bemerkungen über Definition und Messung der Frequenz.* – AEÜ, 2: 167–169.

29. Kock, W. E. (1935): *On the principle of uncertainty in sound.* – J. Acoust. Soc. Amer., 7: 56–58.

30. Konishi, M. (1969): *Time resolution by single auditory neurons in birds.* – Nature, 222: 566–567.

31. Korn, T. S. (1968): *La notation de la fréquence du son.* – Acustica, 20: 55–61.

32. Kriksunóv, V. G. 1965. *Avtomaticheskie analizatory spektrov elektricheskikh signalov (Automatic spectrum analyzers for electric signals).* – Kiev, Technika.

33. Kurze, U. (1965): *Bemerkungen und Versuche zur Unschärferelation beim Ohr.* – 5th Int. Cong. Acoust., Liège, Paper G24.

34. Küpfmüller, K. (1949): *Systemtheorie der elektrischen Nachrichtenübertragung.* – Stuttgart, S. Hirzel.

35. Léon, P. R. & Martin, Ph. (1970): *Prolégomènes à l'étude des structures intonatives.* – Studia Phonetica, 2: 225, particularly „Quatrième Partie".

36. Liang Chin-an & Chistovich, L. A. (1960): *Differentsialnyie porogi po chastote v zavisimosti ot dlitel'nosti tonal'noj posylki.* – Akusticheskii Zhurnal, 6: 81–86. English translation: 1961. Frequency difference limens as a function of tonal duration. Sov. Phys. Acoust., 6: 75–80.

37. Majernik, V. (1964): *Some information-theoretical characteristics of the human ear for non-stationary acoustical stimuli.* – 4th Int. Cong. Cybern. Namur 556–564.

38. Majernik, V. (1967): *Die Unschärferelationen für das Gehör.* – Fyzikálny casopis, 17: 65–72.

39. Marler, P. (1969): *Tonal quality of bird sounds.* – In: Bird Vocalizations, their relation to current problems in biology and psychology. Essays presented to W. H. Thorpe (R. A. Hinde, Ed.), Cambridge University Press: 5–18.

40. McKinney, N. P. (1965): *Laryngeal frequency analysis for linguistic research.* – University of Michigan, Com. Sci. Lab. Rept., No. 14.

41. Nordmark, J. O. (1968): *Mechanisms of frequency discrimination.* – J. Acoust. Soc. Amer., 44: 1533–1540, particularly Sec. I–A.

42. Obata, J. & Kobayashi, R. (1937): *A direct-reading pitch recorder and its application to music and speech.* – J. Acoust. Soc. Amer., 9: 156–161.

43. Obata, J. & Kobayashi, R. (1938): *An apparatus for direct recording the pitch and intensity of sound.* — J. Acoust. Soc. Amer., 10: 147—149.

44. Obata, J. & Kobayashi, R. (1940): *Further applications of our direct-reading pitch and intensity recorder.* — J. Acoust. Soc. Amer., 12: 188—192.

45. Pimonow, L. (1962): *Vibrations en régime transitoire.* — Analyse physique et physiologique, Paris, Dunod.

46. Pumphrey, R. J. (1960—61): *Hearing in birds.* — In: Marshall, A. J.: Biology and comparative anatomy of birds. London, Vol. 2.

47. Righini, G. U. (1964): *Misura della frequenza istantanea della voce.* — Alta Frequenza, 33: 18—21.

48. Ronken, D. A. (1971): *Some effects of bandwidth-duration constraints on frequency discrimination.* — J. Acoust. Soc. Amer., 49: 1232—1242.

49. Schief, R. (1963): *Koinzidenz-Filter als Modell für das menschliche Tonhöhenunterscheidungsvermögen.* — Kybernetik, 2: 8—15.

50. Scharzkopff, J. (1949): *Über Sitz und Leistung von Gehör und Vibrationssinn bei Vögeln.* Zeitschr. vergl. Physiol., 31: 527 (124)

51. Schwartzkopff, J. (1952): *Über den Gehörsinn der Vögel.* — Journal. f. Orn., 93: 91—103.

52. Schwartzkopff, J. (1955): *On the hearing of birds.* — The Auk, 72: 340—7.

53. Sekey, A. (1962): *Auditory adaptation in pitch shift detection.* — 4th Int. Cong. Acoust., Copenhagen, Paper H59.

54. Sekey, A. (1963): *Short-term auditory frequency discrimination.* — J. Acoust. Soc. Amer., 35: 682—690.

55. Soanes, S. V. (1952): *Some problems in audio frequency spectrum analysis.* — Electron. Eng., 24, Part 1, 268—270.

56. Stewart, G. W. (1931): *Problems suggested by an uncertainty principle in acoustics.* — J. Acoust Soc. Amer., 2: 325—329.

57. Stopskij, S. B. (1962): *Analizatory spektra zvukorykh i infrazvukorykh chastot dl'a akusticheskoi spektrometrii. (Sound and infrasound spectrum analyzers for acoustic spectrometry.)* — Leningrad, Gosenergoizdat.

58. Sundberg, J. & Tjerlund P. (1971): *Real time notation of performed melodies by means of a computer.* — 7th Int. Cong. Acoust., Budapest, Paper 21, S6.

59. Szőke, P. (1974): *General concept and the three specific evolutionary stages of music: on the level of 1. pure physical (pre-biological) existence, 2. animal life, and 3. human society.* — Biophon, III, 1: 5—10. (IBAC, Aarhus.)

60. Szőke, P., Gunn, W. W. H. & Filip, M. (1969): *The musical microcosm of the Hermit Thrush.* — Studia Musicologica, 11: 423—438.

61. Szőke, P., Tarnóczy, T. & Filip, M. (1970): *Inadequacy of sound spectrography in certain bioacoustical applications.* — Lecture presented at the 15th Int. Ornithol. Cong., The Hague.

62. Thielcke, G. (1966): *Die Auswertung von Vogelstimmen nach Tonbandaufnahmen.* — Die Vogelwelt, 87, Heft 1; 1—14.

63. Thorpe, W. H. (1961): *Bird-Song. The biology of vocal communication and expression in birds.* — Cambridge University Press: 1—143.

64. Thorpe, W. H., Lade B. I. (1961*a*): *The songs of some families of the Passeriformes. I. Introduction: The analysis of bird songs and their expression in graphic notation.* — The Ibis, 103a: 231—245.

65. Tove, P. A., Norman, B., Isaksson, L. & Czekajewski, J. (1966): *Direct-recording frequency and amplitude meter for analysis of musical and other sonic waveforms.* — J. Acoust. Soc. Amer., 39: 362—371.

66. Tove, P. A., Isaksson, L., Normann, B. & Czekajewski, J. (1966*a*): *Frekvensmätare med kort inställningstid.* — Elektronik i teori och praktik, No. 4: 64—73.

67. Tove, P. A. & Pelc, K. (1964): *Spectrum analysis of sonic signals.* Elteknik, 7: 19 – 26.

68. Ungeheuer, G. (1963): *Zur Periodizitätsanalyse phonetischer Signale.* – Phonetica, 10: 174 – 186.

69. Winckel, F. (1960): *Phänomene des musikalischen Hörens.* – Berlin, Max Hesse.

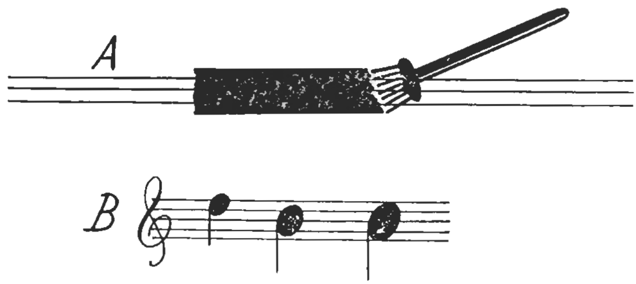


Fig. 1. The frequency (pitch) smearing effect of wide bands

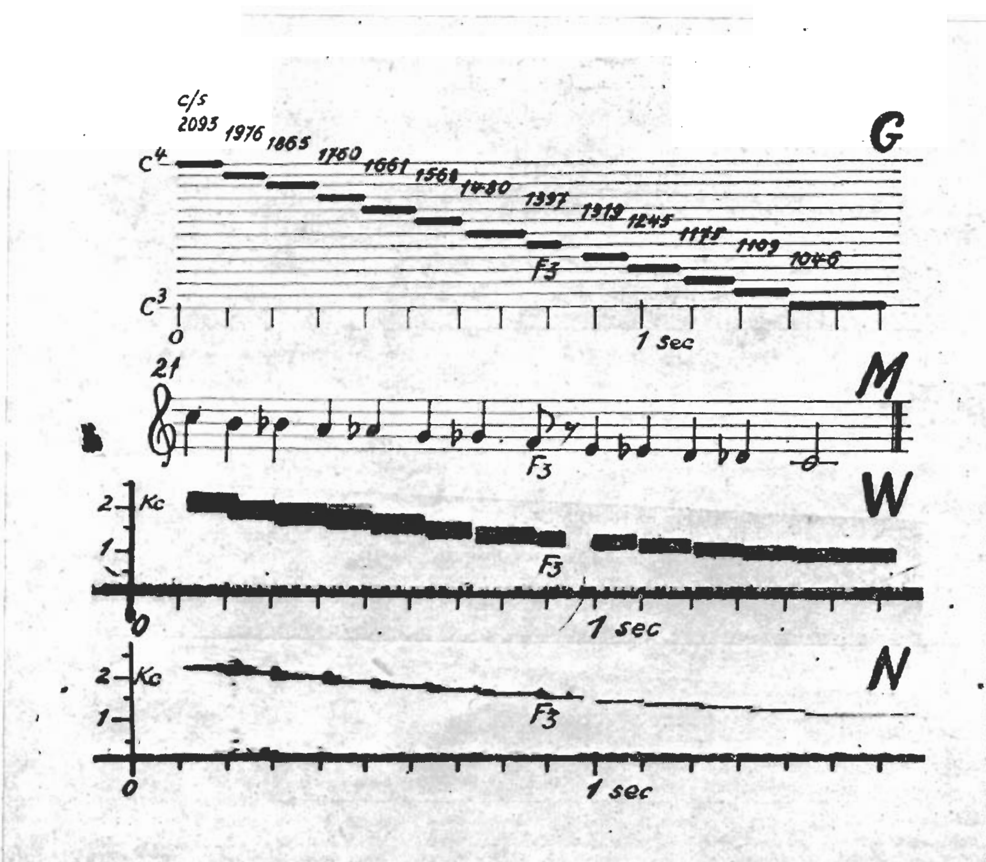


Fig. 2. Different illustrations of the semitone scale C_4 to C_3

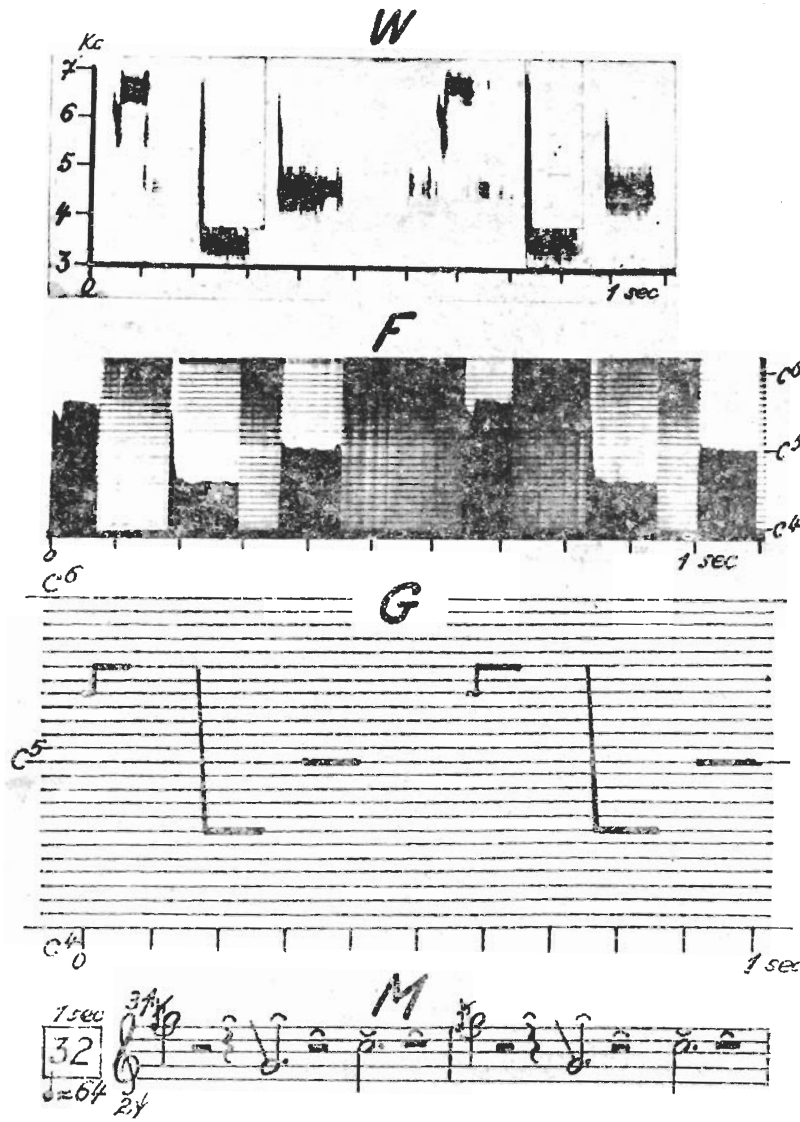


Fig. 3. Great Tit song portion

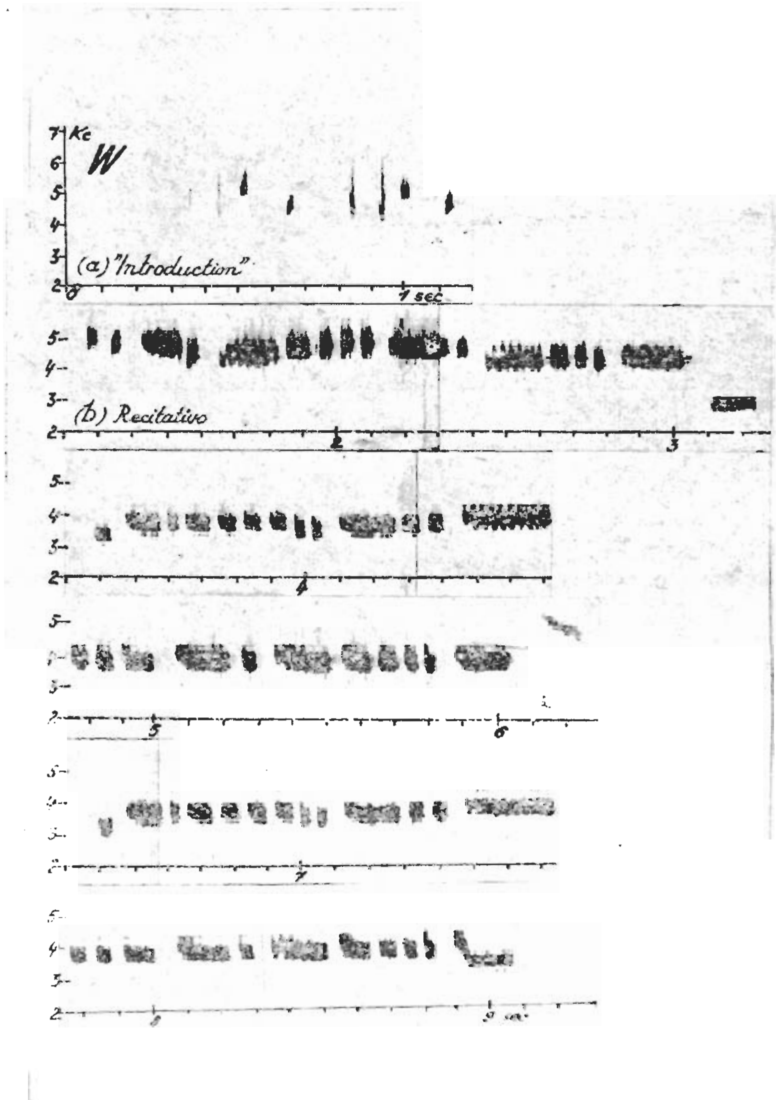


Fig. 4. Grey Warbler song

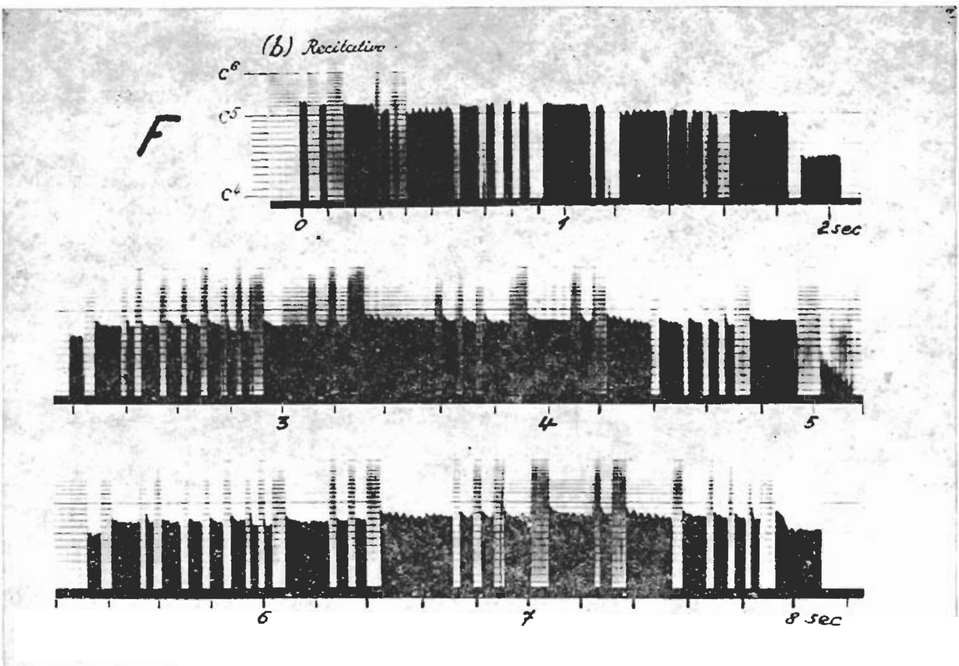


Fig. 4F. Grey Warbler, instantaneous frequency graph of Fig. 4W

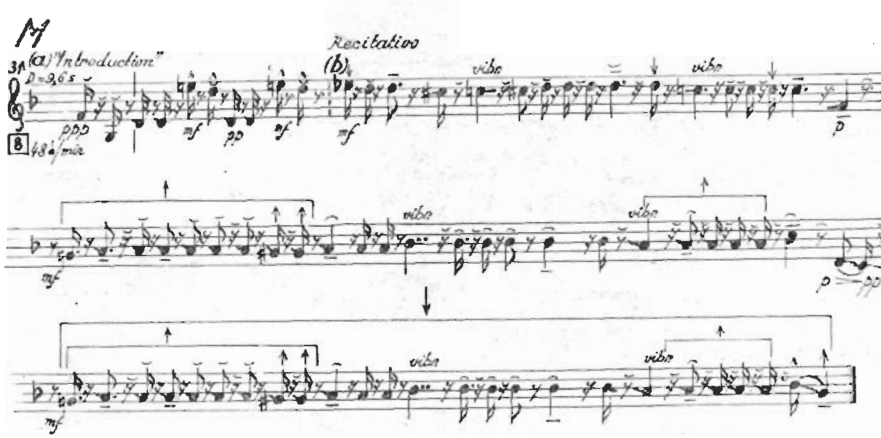


Fig. 4M. Grey Warbler, musical representation of Fig. 4 F(W)

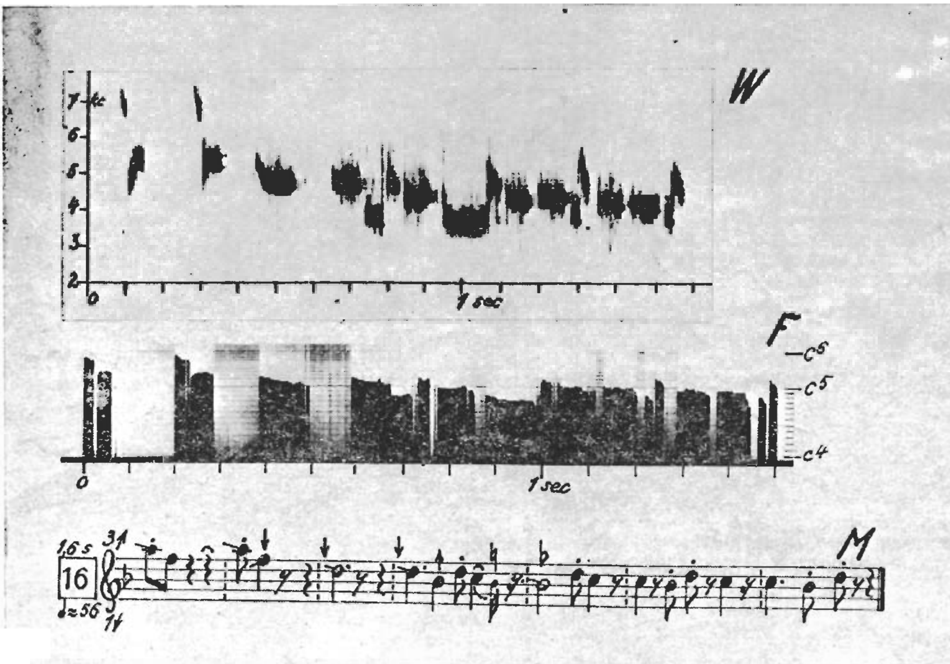


Fig. 5. Yellow-breasted Tit song

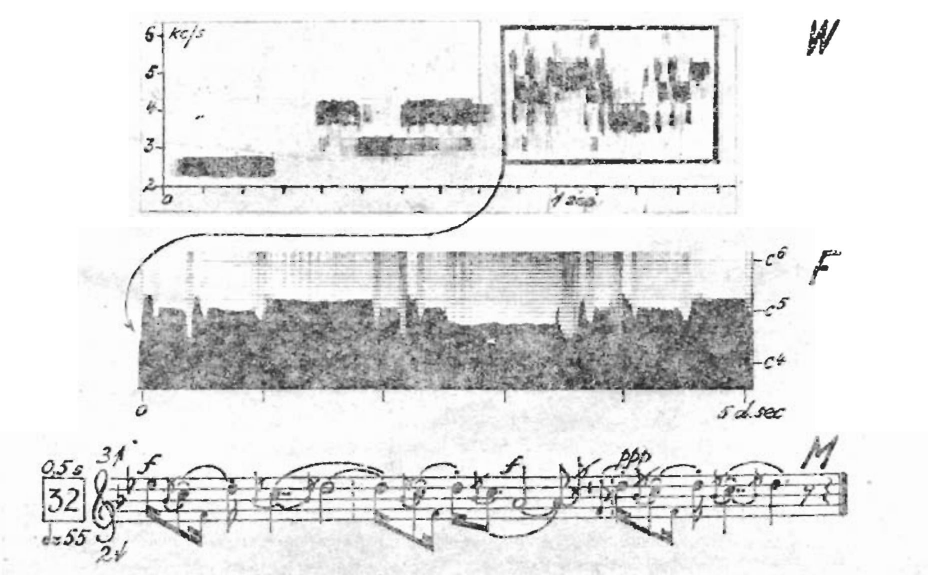


Fig. 6. Hermit Thrush song

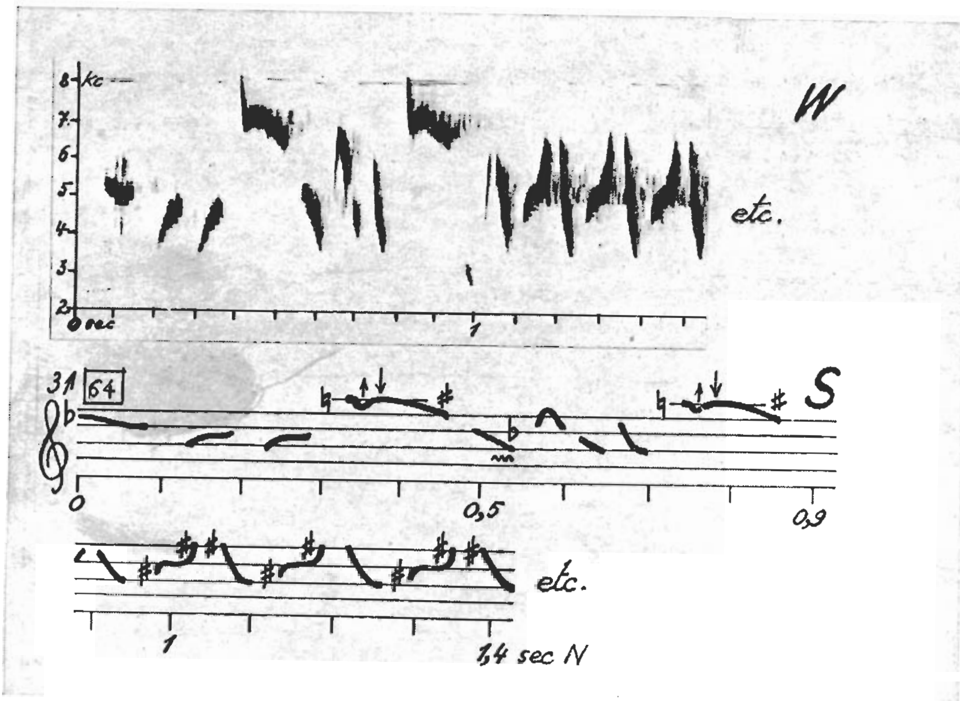


Fig. 7. Wren, non-musical song (initial part)

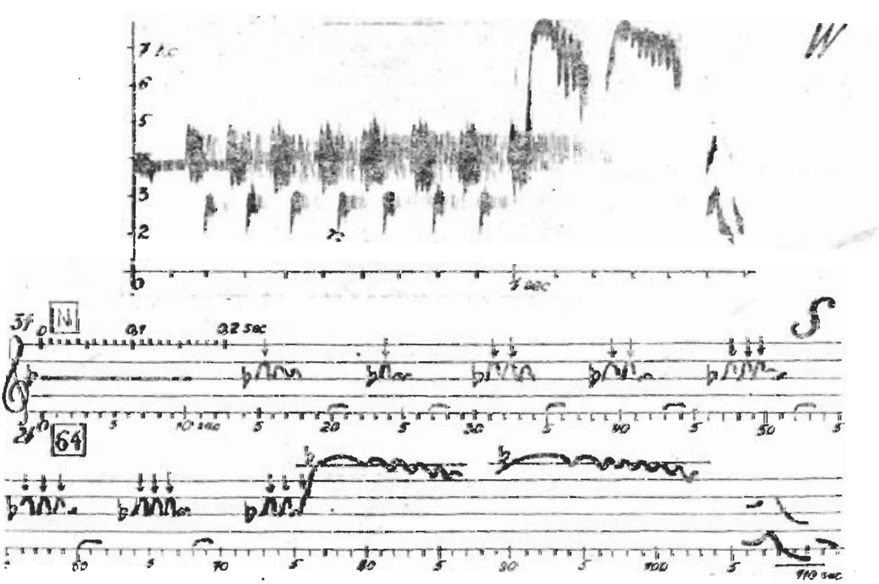


Fig. 8 Redstart, non-musical song

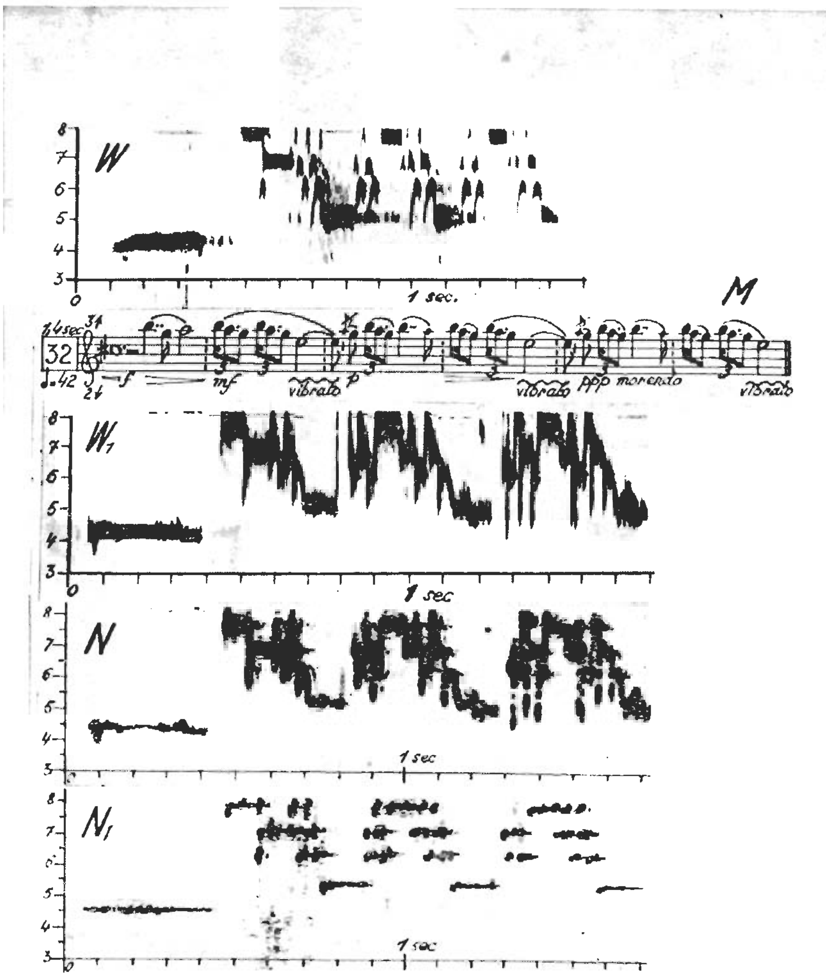


Fig. 9. Hermit Thrush, a folksong-like micro-melody

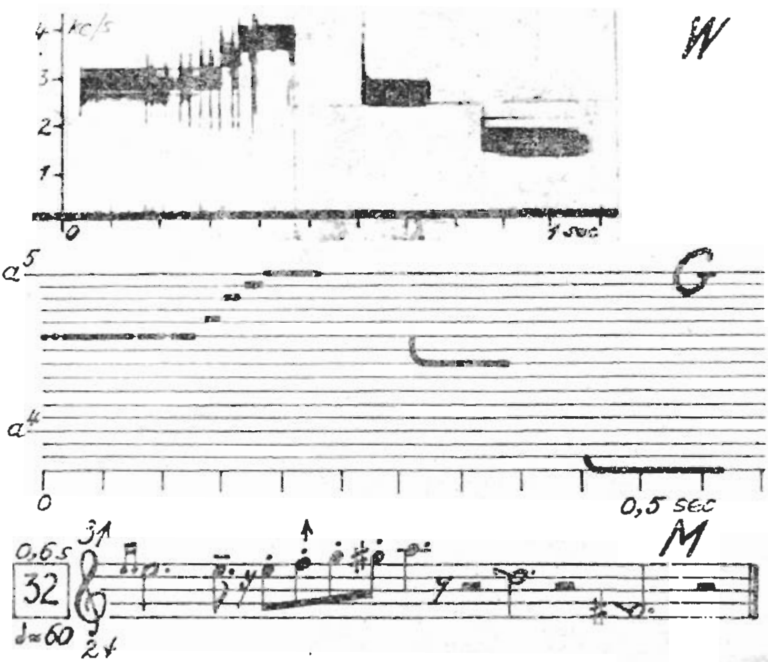


Fig. 10. Great Tit call of musical structure

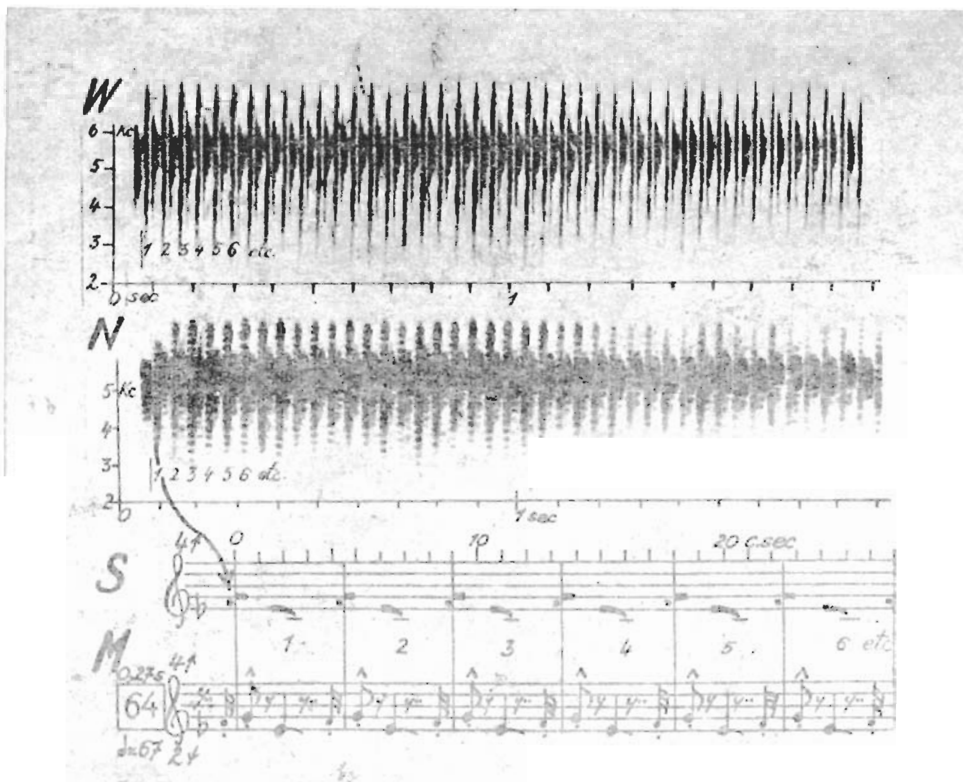


Fig. 11. Grasshopper Warbler, song portion

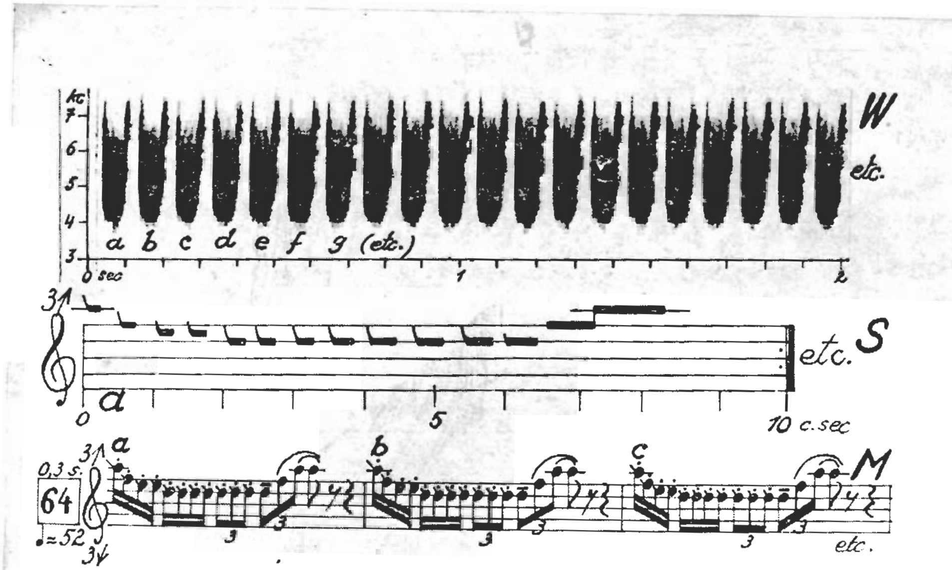


Fig. 14. River Warbler, song portion

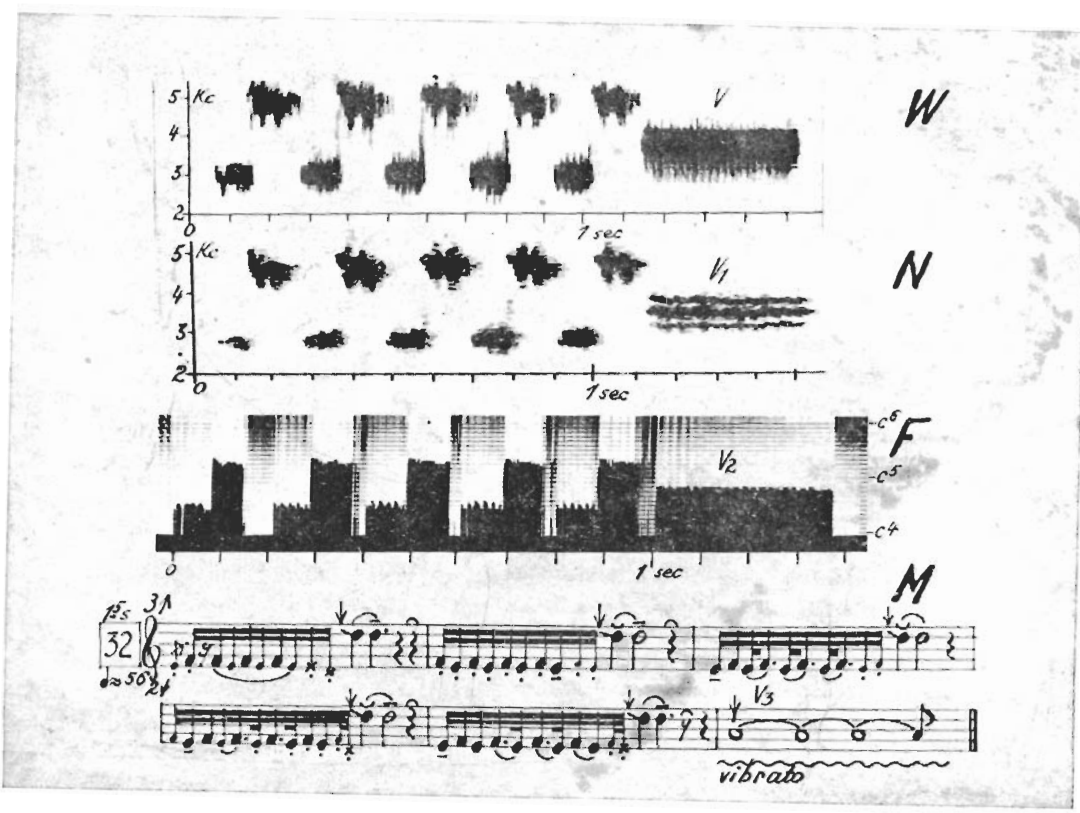


Fig. 15. Ortolan Bunting song

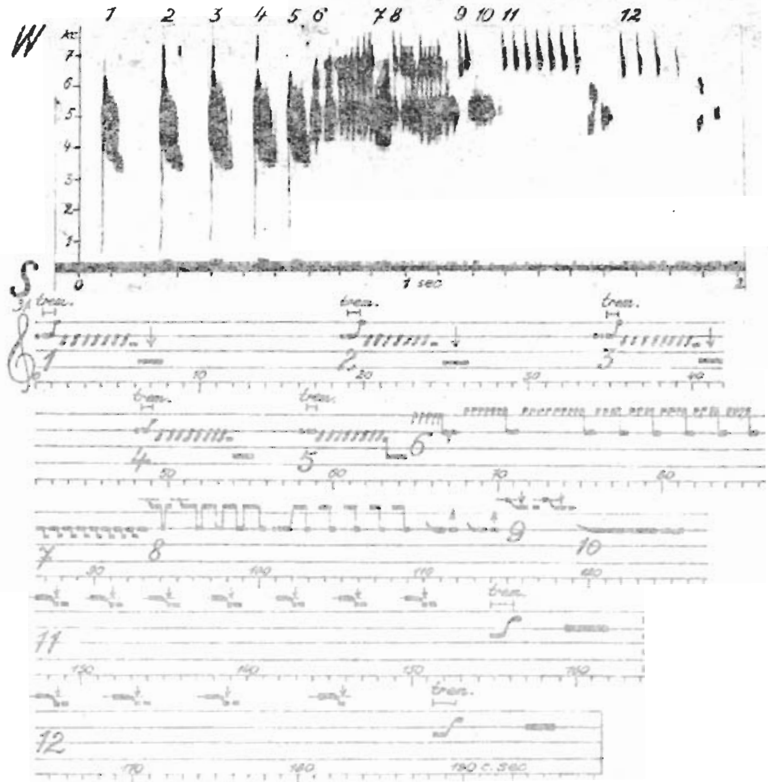


Fig. 16. Corn Bunting song

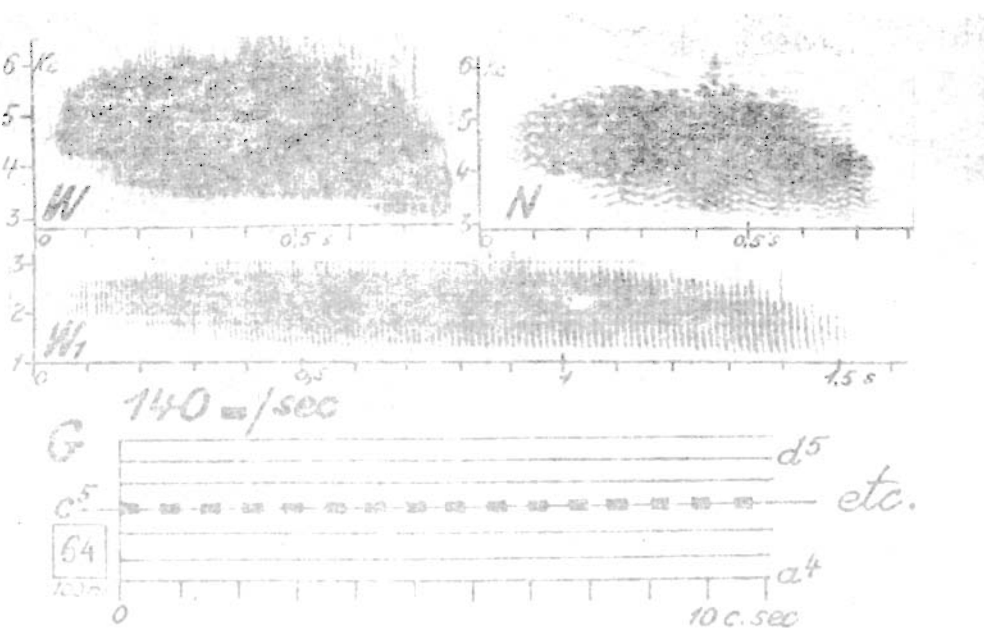


Fig. 17. Greenfinch call

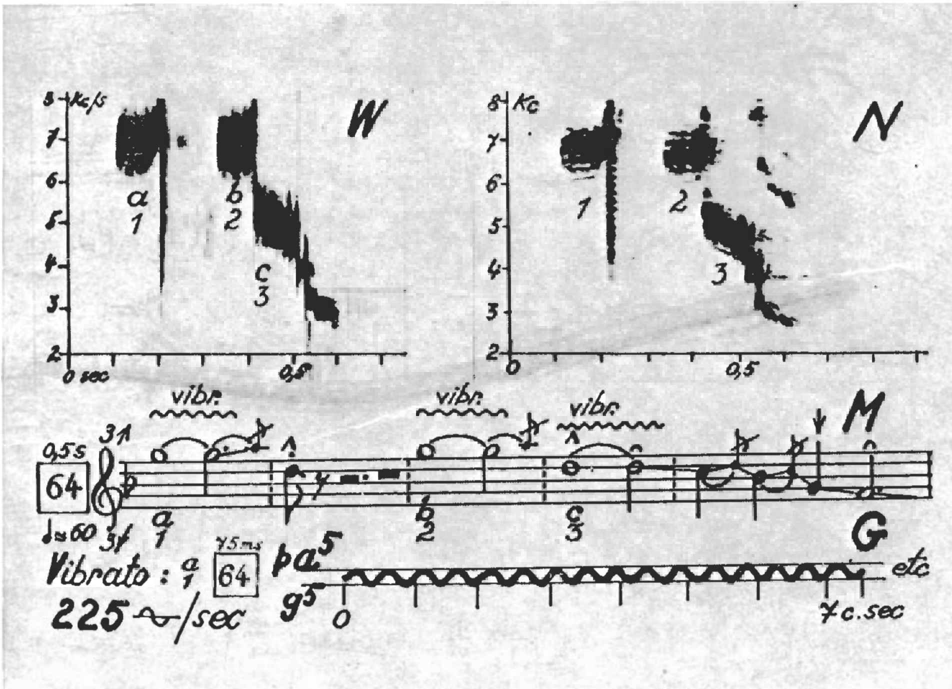


Fig. 18. Great Tit, musically structured call with vibratos